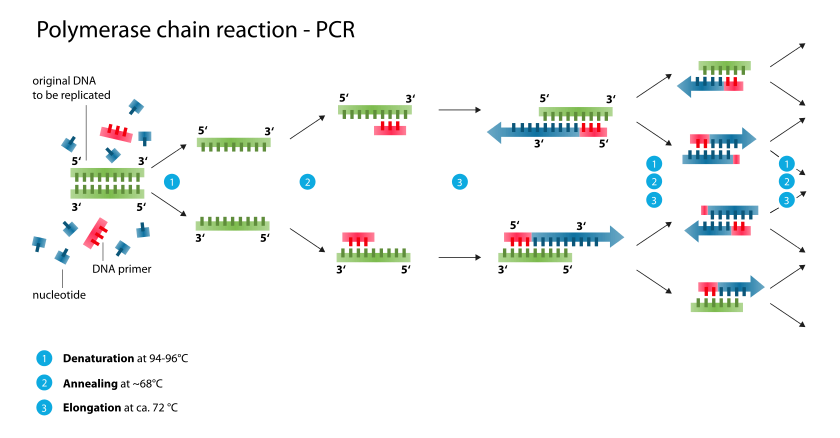
There are two main types of primers that people use:

1. Those that amplify DNA
2. Those that modify DNA

We will focus on the first. We will use the ApoE protein as a model. Mutations in it are a strong genetic marker for Alzheimer’s disease. The mutations in this gene occur at amino acids 112 and 158. Our goal will be to amplify this gene and send it out for sequencing. First we will create primers for the whole gene and then just the important region.

First, find the sequence of the gene that you want to amplify or modify. A great place to look is NCBI ([http://www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/nuccore/DQ286969.1)). I searched and found the Human ApoE mRNA gene sequence (<http://www.ncbi.nlm.nih.gov/nuccore/FJ525876.1>). There are about 3.2 Billion bases of DNA in human beings so that means our primers should probably be at least that complex. Because there are 4 bases in DNA the complexity is equal to 4 to the power of the length. A DNA of 16 bases pairs will occur about 1 in 4.3 Billion times(416). Generally, primers are between 15 and 21 base pairs in length.

First we need to understand how DNA is copied. DNA is amplified from the 3’(pronounced Three Prime) end to the 5’(pronounced Five Prime) end of the strand being copied. The amplification is said to go in the 5’ to 3’ direction.



The forward primer is easy and is the primer that resides on the bottom strand on the 3’ side. The reverse primer is more complicated and binds to the top strand on the 3’ side.

Let’s first do an example

Here is our DNA:

5’ - ATGAGGTTTGCAGCG - 3’

| | | | | | | | | | | | | | |

3’ - TACTCCAAACGTCGC - 5’

let’s make hypothetical primers for the short pieces of DNA that are 4 bases each.

The forward primers need to bind to the 3’ end of the bottom strand and so is identical to the top strand! That means our hypothetical forward primer would be ATGA. Because primers are read and created by humans our reverse primer need to be written from the beginning to the end. This is called the “reverse complement” of the top strand. The 4 bases that bind to the 3’ of the top strand are TCGC. But remember that the primer starts at the 3’ end so it should be read as CGCT. This is the reverse complement, the reverse of the opposite of the top strand.

Looking at the sequence of ApoE try and make 20 base forward and reverse primers. The answers are below. <http://www.ncbi.nlm.nih.gov/nuccore/FJ525876.1>

ApoE forward Primer: cagcggatccttgatgctgc

ApoE Reverse Primer: aagcaccaagttcagggtgt

We can use these primers to amplify DNA extracted from yourself. Clean it up and send it off for sequencing. But….. sequencing reads even good ones are generally only about 1000 bases in length and the gene is > 7000 bases.

Primers don’t need to be created just at the end of DNA you can create them anywhere to amplify the DNA. Now Let’s amplify the region around 112 to 158 so we can have a more reasonable sequence.

First let’s translate the DNA to find out where these amino acids are located in the DNA using <http://web.expasy.org/translate/> Select “Include Nucleotide Sequence”. Now let’s search for a string of amino acids that are found in the translation on the NCBI page “VCG”. Because of the spaces in the translate webpage we need to search on the page for “V C G”. It should be in the 5’ Frame 2.

The bases in a primer do not show up in the sequence and usually the first few bases noisy so let’s create a primer that is about 50-100 bases upstream of the VCG.

I selected as my forward primer: gagacgcgggcacggctgtc

The reverse primer needs to be about 50-100 bases downstream of R158. So let’s find the DNA that is associated with the VRL sequence, which are amino acids 157-159. I chose this sequence: agcgcctggcagtgtaccag

but that isn’t the primer we need to make the reverse complement.

The complement is: tcgcggaccgtcacatggtc

The reverse complement is: ctggtacactgccaggcgct

The forward and reverse primers to amplify the ApoE protein to see if we contain mutations that are a predictor of Alzheimer’s:

Forward: gagacgcgggcacggctgtc

Reverse: ctggtacactgccaggcgct

Finally, you want to search the human genome for your primers to make sure that the sequence is not duplicated any place else (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>)

* Use one primer at a time
* Choose Human Genomic + Transcript for database
* Click BLAST

Looks like the closest is 75% identity, which is kind of bad but not awful.

One can order primers from IDT: <http://www.idtdna.com/site>